Spread of a contagion in a heterogeneous population- Linear algebra tutorial

Yuval Peres

August 2020

1 Geometric growth

In the simplest model, suppose that every infected person meets m people per day (on average) and each such meeting yields a virus transmission with probability p. Moreover, suppose every infected person recovers (or is otherwise removed from the infectious set) with probability α each day. In a population of size N, let

$$I_t = \{\text{The number of infected individuals on day } t\}$$

Then ignoring random fluctuations and denoting $\beta = pm$,

$$I_{t+1} = I_t + \beta I_t - \alpha I_t = I_t(1+\gamma) \tag{1}$$

where $\gamma = \beta - \alpha$.

Despite its extreme simplicity, this formula often represents initial growth of epidemics quite well. The factor $1+\gamma$ is often denoted r_0. If $\gamma > 0$ the epidemic grows exponentially, while if $\gamma < 0$ it decays.

2 Heterogeneous populations

Suppose we consider two communities. In the first the contagion has spread widely and is now shrinking due to effective mitigation (e.g. lockdown), with $I_0 = 10^4$, $I_{t+1} = 0.9 \cdot I_t$. In the second community, the contagion is much smaller but growing rapidly: $I_0 = 10$, $I_{t+1} = 2I_t$.

If these two communities are considered together, the situation may look rosy in the first few days, but can get much worse later. A plot with these parameters can be found in Figure 2; note the total number of people infected decreases in the first six days but starts increasing afterwards.

A more complex situation arises when multiple types of individuals interact. For a concrete example, suppose that a fraction θ of the population are cashiers and the rest are customers.

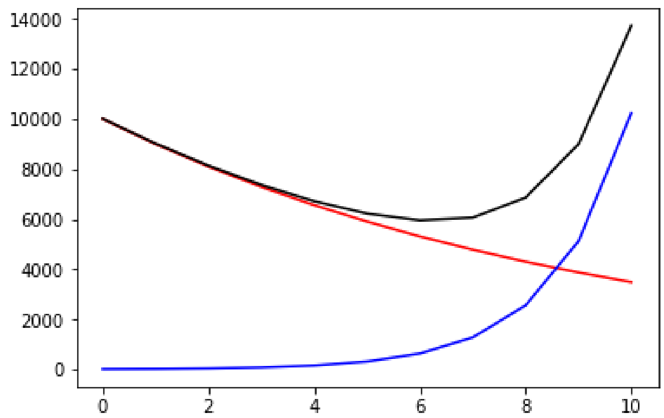

Figure 1: The X axis denotes the time and the Y axis the number of people infected each day, where red shows the people in the first community, blue the people in the second community, and black the sum.

Every cashier interacts, on average, with m_1 customers and with one other cashier per day. Every customer interacts, on average, with m_2 cashiers and one other customer per day.

Naturally, $m_1 > 1 > m_2$. Moreover, counting the expected number of customer-cashier interactions per day shows that

$$\theta m_1 = (1 - \theta) m_2 \tag{2}$$

Each interaction between an infected person and a susceptible one yields transmission with probability p. For specific values, suppose that

$$\theta = 1/161, m_1 = 40, m_2 = 1/4, \text{ and } p = 1/10, \tag{3}$$

so (2) holds. We also assume that every infected person is removed (recovered or quarantined) with probability α (say with $\alpha = 1/5$).

Denote by I_t the number of infected customers and by J_t the number of infected cashiers. Suppose initially that $I_0 = 100$ and $J_0 = 0$. The next plot shows the numbers I_t and J_t over a period of 10 days.

In the early phases of the epidemic, the number of susceptibles hardly changes. So on average,

$$\begin{aligned} I_{t+1} &= I_t + p \cdot \big(I_t + m_1 J_t\big) - \alpha I_t \\ J_{t+1} &= J_t + p \cdot \big(J_t + m_2 I_t\big) - \alpha J_t \end{aligned} \tag{4}$$

How does this contagion spread? Consider the numerical example in (3), where $\alpha = 1/5$. During the five days (on average) that a customer is contagious, they will interact (on average) with $25/4$ people and infect in expectation $25p/4 = 5/8$ people. For a cashier, the corresponding number is $101/2$.

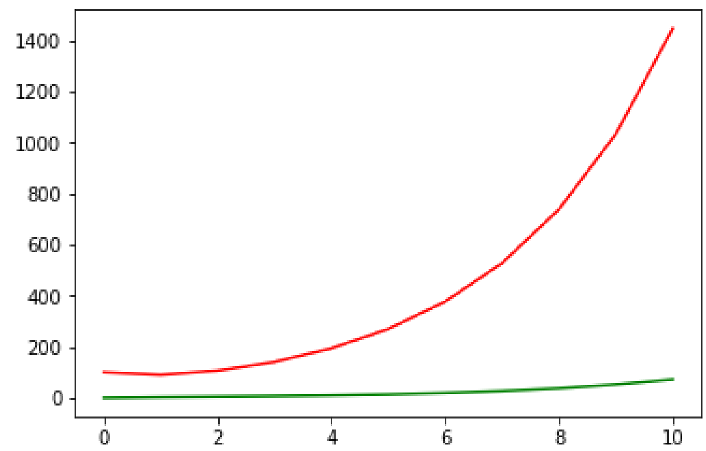

Figure 2: The X axis denotes the time and the Y axis the number of people infected each day, where red shows the number of infected customers (I_t) and blue the number of infected cashiers (J_t).

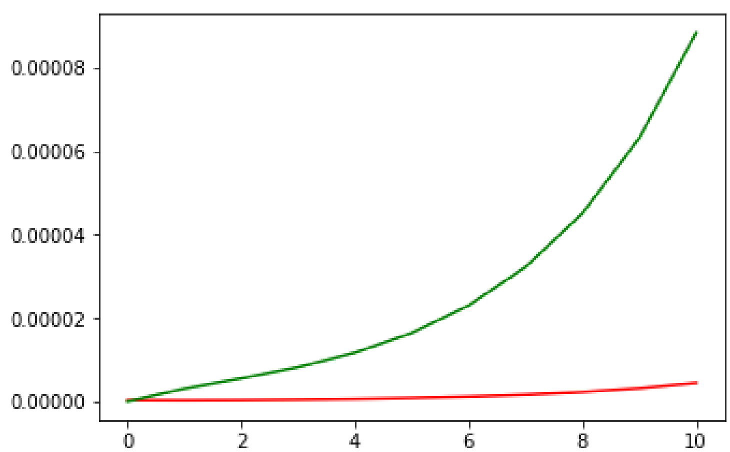

Figure 3: The X axis denotes the time and the Y axis the risk in the early stage of the epidemic for customers (in red) and for cashiers (in green).

Since cashiers comprise only $\theta = 1/401$ of the population, naive averaging would yield that the expected number of people an infected person would infect is

$$\frac{400}{401} \cdot \frac{5}{8} + \frac{1}{401} \cdot \frac{101}{2} < 0.75,$$

which suggests the contagion is shrinking.

However, this heuristic is faulty. Rewriting (4) in matrix form, we have

$$\begin{pmatrix} I_{t+1} \\ J_{t+1} \end{pmatrix} = A \begin{pmatrix} I_t \\ J_t \end{pmatrix}$$

where

$$A = \begin{pmatrix} 9/10 & 10 \\ 1/40 & 9/10 \end{pmatrix} \tag{5}$$

We deduce that

$$\begin{pmatrix} I_t \\ J_t \end{pmatrix} = A^t \begin{pmatrix} I_0 \\ J_0 \end{pmatrix}$$

To compute the powers of A^t we diagonalize A. First find eigenvalues and eigenvectors by solving $Av = \lambda v$, that is

$$\begin{cases} \frac{9v_1}{10} + 10v_2 = \lambda v_1 \\ \\ \frac{v_1}{40} + \frac{9v_2}{10} = \lambda v_2 \end{cases}$$

Equivalently,

$$\begin{cases} 10v_2 = \left(\lambda - \frac{9}{10}\right)v_1 \\ \\ \frac{v_1}{40} = (\lambda - \frac{9}{10})v_2 \end{cases}$$

Multiplying these equations gives that $1/4 = (\lambda - 9/10)^2$, so $\lambda \in \{7/5, 2/5\}$.

For $\lambda = 7/5$ we get $v = \begin{pmatrix} 20 \\ 1 \end{pmatrix}$; for $\tilde{\lambda} = 2/5$, we get $\tilde{v} = \begin{pmatrix} -20 \\ 1 \end{pmatrix}$.

Represent

$$\begin{pmatrix} I_0 \\ J_0 \end{pmatrix} = \begin{pmatrix} 100 \\ 0 \end{pmatrix} = \frac{5}{2}(v - \tilde{v})$$

Thus

$$\begin{pmatrix} I_t \\ J_t \end{pmatrix} = \frac{5}{2}A^t(v - \tilde{v}) = \frac{5}{2}\Big[\left(\frac{7}{5}\right)^t v - \left(\frac{2}{5}\right)^t \tilde{v}\Big] \tag{6}$$

The first summand grows exponentially in t, while the second shrinks. So the infection grows with growth factor $7/5$.

3 Introducing the SIR model

In a population of size N, write

$$\begin{aligned} S_t &= \{\text{The number of susceptible individuals on day } t\} \\ I_t &= \{\text{The number of infected individuals on day } t\} \\ R_t &= \{\text{The number of removed individuals on day } t\} \end{aligned}$$

Here, removed can indicate recovery or showing symptoms that lead to quarantine, so an individual cannot infect others.

Suppose each infected person meets m people each day, on average, and a fraction p of these meetings result in transmission of the infection. Moreover, each infected person is removed with probability α each day. Then, with $\beta = mp$, we have

$$\begin{aligned} S_{t+1} &= S_t - \beta \frac{S_t}{N} I_t \\ I_{t+1} &= I_t + \beta \frac{S_t}{N} I_t - \alpha I_t \\ R_{t+1} &= R_t + \alpha I_t \end{aligned} \tag{7}$$

This is the celebrated SIR model [1].

In the early stages of the infection, S_t/N is close to 1, so equation (7) for I_{t+1} is close to the geometric sequence $I_{t+1} = (1+\gamma)I_t$, where $\gamma = \beta - \alpha$.

References

[1] Herbert W. Hethcote. The mathematics of infectious diseases. *SIAM Review*, 42(4):599–653, 2000.